EE793 Target Tracking: Lecture 1
Introduction to State Estimation

Umut Orguner
umut@metu.edu.tr
room: EZ-12 tel: 4425

Department of Electrical & Electronics Engineering
Middle East Technical University
Ankara, Turkey

Outline

Bayesian State Estimation

Problem Definition

Consider the state-space system

$$x_{k+1} = f(x_k) + w_k$$
$$y_k = h(x_k) + v_k$$

where

- $x_k \in \mathbb{R}^{n_x}$ is the state with the initial state $x_0 \sim p(x_0)$;
- $y_k \in \mathbb{R}^{n_y}$ is the measurement;
- $w_k \in \mathbb{R}^{n_x}$ is the white process noise with a known distribution $p(w)$ independent from x_k;
- $v_k \in \mathbb{R}^{n_y}$ is the white measurement noise with a known distribution $p(v)$ independent from x_k.

Aim: Find the posterior density of the state $p(x_k|y_{1:k})$ where

$$y_{1:k} \triangleq \{y_1, y_1, \ldots, y_k\}.$$

Bayesian State Estimation

- The process noise represents our lack of knowledge about the system dynamics. The larger the process noise, the smaller will be our trust on the state equation.
- The measurement noise represents the imperfections in acquiring the data. The larger the measurement noise, the smaller will be our trust on the measurements.
- Bayesian state estimation, except for few special cases, boils down to an infinite dimensional estimation problem, i.e., a function ($p(x_k|y_{1:k})$) has to be computed.
- Basic probability theory gives a recursive solution in the form

$$p(x_{k-1}|y_{1:k-1}) \xrightarrow{\text{prediction}} p(x_k|y_{1:k-1}) \xrightarrow{\text{update}} p(x_k|y_{1:k})$$

Solution: Bayesian Density Recursion

Bayesian Recursion

- Start with $p(x_0)$, set $k = 1$.
- For each k
  - Prediction Update

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})\,\mathrm{d}x_{k-1}$$

  - Measurement Update

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})}$$

    where

$$p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})\,\mathrm{d}x_k$$

    is constant with respect to x_k.
  - $k = k + 1$.

Solution: Bayesian Density Recursion

Terminology

- $p(x_k|y_{1:k-1})$: Predicted state density
- $p(y_k|y_{1:k-1})$: Predicted measurement density
- $p(x_k|y_{1:k})$: Estimated state density/ posterior state density
- $p(y_k|x_k)$: Measurement likelihood
- $p(x_k|x_{k-1})$: State transition density

Point Estimates

MMSE criterion

- Define the estimates as

$$\hat{x}^{\text{MMSE}}_{k|k-1} = \arg\min_{\hat{x}_k} E\left[\|x_k - \hat{x}_k\|_2^2|y_{1:k-1}\right]$$
$$\hat{x}^{\text{MMSE}}_{k|k} = \arg\min_{\hat{x}_k} E\left[\|x_k - \hat{x}_k\|_2^2|y_{1:k}\right]$$

  which minimize the mean square (estimation or prediction) error.
- The estimates are given as

$$\hat{x}^{\text{MMSE}}_{k|k-1} = E\left[x_k|y_{1:k-1}\right]$$
$$\hat{x}^{\text{MMSE}}_{k|k} = E\left[x_k|y_{1:k}\right]$$

  which are the means for the predicted and estimated state densities.

Point Estimates

- MMSE is the most common criterion to obtain point estimates.
- The second common point estimate is called maximum a posteriori (MAP) estimate.

MAP criterion

The estimates are given as

$$\hat{x}^{\text{MAP}}_{k|k-1} = \arg\max_{x_k} p(x_k|y_{1:k-1})$$
$$\hat{x}^{\text{MAP}}_{k|k} = \arg\max_{x_k} p(x_k|y_{1:k})$$

which are the global maxima for the predicted and estimated state densities.

Point Estimates

Uncertainty Measures

Every point estimate must be accompanied by an uncertainty measure describing how trustable it is.

The most common uncertainty measure is the covariance.

$$P_{k|k-1} = E\left[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}} | y_{1:k-1}\right]$$
$$P_{k|k} = E\left[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{\mathrm{T}} | y_{1:k}\right]$$

which are the covariances of the prediction $\hat{x}_{k|k-1}$ and the estimate $\hat{x}_{k|k}$.

Most Important Special Case

Original Problem

$$x_{k+1} = f(x_k) + w_k$$
$$y_k = h(x_k) + v_k$$

with $w_k \sim p(w_k)$, $v_k \sim p(v_k)$ and $x_0 \sim p(x_0)$.

Special Case: Linear Gaussian Systems

- $f(x_k) = Ax_k$ where $A \in \mathbb{R}^{n_x \times n_x}$;
- $g(x_k) = Cx_k$ where $C \in \mathbb{R}^{n_y \times n_x}$;
- $w_k \sim \mathcal{N}(w_k; 0, Q)$ where $Q \geq 0 \in \mathbb{R}^{n_x \times n_x}$;
- $v_k \sim \mathcal{N}(v_k; 0, R)$ where $R > 0 \in \mathbb{R}^{n_y \times n_y}$;
- $x_0 \sim \mathcal{N}(x_0; \hat{x}_{0|0}, P_{0|0})$.

Linear Gaussian Systems

Special Problem

$$x_{k+1} = Ax_k + w_k$$
$$y_k = Cx_k + v_k$$

with $w_k \sim \mathcal{N}(w_k; 0, Q)$, $v_k \sim \mathcal{N}(v_k; 0, R)$ and $x_0 \sim \mathcal{N}(x_0; \hat{x}_{0|0}, P_{0|0})$.

In this case it can be shown that all densities are Gaussian:

- $p(x_k|y_{1:k-1}) = \mathcal{N}(x_k; \hat{x}_{k|k-1}, P_{k|k-1})$
- $p(y_k|y_{1:k-1}) = \mathcal{N}(y_k; \hat{y}_{k|k-1}, S_{k|k-1})$
- $p(x_k|y_{1:k}) = \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k})$
- $p(y_k|x_k) = \mathcal{N}(y_k; Cx_k, R)$
- $p(x_k|x_{k-1}) = \mathcal{N}(x_k; Ax_{k-1}, Q)$

Linear Gaussian Systems

- Since the density $p(x_k|y_{1:k})$ is always Gaussian, it is possible to keep only its sufficient statistics $\hat{x}_{k|k}$ and $P_{k|k}$
- In other words, instead of propagating densities as

$$p(x_{k-1}|y_{1:k-1}) \xrightarrow{\text{prediction}} p(x_k|y_{1:k-1}) \xrightarrow{\text{update}} p(x_k|y_{1:k})$$

we propagate only the means and the covariances as

$$\hat{x}_{k-1|k-1}, P_{k-1|k-1} \xrightarrow{\text{prediction}} \hat{x}_{k|k-1}, P_{k|k-1} \xrightarrow{\text{update}} \hat{x}_{k|k}, P_{k|k}.$$

- As a result, the infinite dimensional estimation problem reduces to a finite dimensional estimation problem.

Kalman Filter

The equations of propagation for the means and the covariances are called Kalman filter.

Kalman Filter

- Start with $\hat{x}_{0|0}$, $P_{0|0}$, set $k = 1$.
- For each k:
  - Prediction Update
$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1}$$
$$P_{k|k-1} = AP_{k-1|k-1}A^{\mathrm{T}} + Q$$
  - Measurement Update
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1})$$
$$P_{k|k} = P_{k|k-1} - K_k S_{k|k-1} K_k^{\mathrm{T}}$$

  where
$$\hat{y}_{k|k-1} = C\hat{x}_{k|k-1}$$
$$S_{k|k-1} = CP_{k|k-1}C^{\mathrm{T}} + R$$
$$K_k = P_{k|k-1}C^{\mathrm{T}}S_{k|k-1}^{-1}$$

Kalman Filter

Terminology

- $\hat{x}_{k|k-1}$: Predicted state
- $P_{k|k-1}$: Covariance of the predicted state
- $\hat{x}_{k|k}$: Estimated state
- $P_{k|k}$: Covariance of the estimated state
- $\hat{y}_{k|k-1}$: Predicted measurement
- $\nu_k \triangleq y_k - \hat{y}_{k|k-1}$: Measurement prediction error / innovation
- $S_{k|k-1}$: Covariance of the predicted measurements / innovation covariance
- K_k: Kalman gain

Nonlinear Non-Gaussian Systems

Original Problem

$$x_{k+1} = f(x_k) + w_k$$
$$y_k = h(x_k) + v_k$$

with $w_k \sim p(w_k)$, $v_k \sim p(v_k)$ and $x_0 \sim p(x_0)$.

- In general assuming that the functions $f(\cdot)$ and $g(\cdot)$ are linear is far too restrictive.
- Similarly the noise terms cannot be assumed to be Gaussian in many cases.
- The exact posterior density $p(x_k|y_{1:k})$ is no longer Gaussian for the general case.

Nonlinear Non-Gaussian Systems

- We are going to consider two main type of solutions to the Bayesian state estimation problem for nonlinear non-Gaussian systems.
- The posterior $p(x_k|y_{1:k})$ can be approximated in two different ways:

$$p(x_k|y_{1:k}) \approx \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}) \quad \text{Gaussian Approximation}$$
$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{N} \pi_{k|k}^{(i)} \delta_{x_{k|k}^{(i)}}(x_k) \quad \text{Particle Approximation}$$

where $\pi_{k|k}^{(i)} \geq 0$ and $\sum_{i=1}^{N} \pi_{k|k}^{(i)} = 1$.

Nonlinear Transformations of Gaussian Random Variables

Main Task

Consider a random vector $\phi \sim \mathcal{N}(\phi; \bar{\phi}, \Phi)$. Let ψ be another random vector related to ϕ as

$$\psi = g(\phi)$$

where $g(\cdot)$ is a nonlinear function. Suppose that we would like to approximate the density $p(\psi)$ of ψ as a Gaussian as follows.

$$p(\psi) \approx \mathcal{N}(\psi; \bar{\psi}, \Psi).$$

Find $\bar{\psi}$ and Ψ.

Linearization

- The first and the most basic solution to this problem is linearization.
- Let us linearize $g(\phi)$ around the mean $\bar{\phi}$.

$$g(\phi) \approx g(\bar{\phi}) + G(\phi - \bar{\phi})$$

where

$$G = \left.\frac{\partial g}{\partial \phi}\right|_{\phi=\bar{\phi}} = \begin{bmatrix} \left.\frac{\partial g_1}{\partial \phi_1}\right|_{\phi=\bar{\phi}} & \cdots & \left.\frac{\partial g_1}{\partial \phi_{n_\phi}}\right|_{\phi=\bar{\phi}} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial g_{n_\psi}}{\partial \phi_1}\right|_{\phi=\bar{\phi}} & \cdots & \left.\frac{\partial g_{n_\psi}}{\partial \phi_{n_\phi}}\right|_{\phi=\bar{\phi}} \end{bmatrix}$$

is the Jacobian.

Linearization

Results from Linearization

The following simple mean and covariance are obtained for $p(\psi) \approx \mathcal{N}(\psi; \bar{\psi}, \Psi)$:

$$\bar{\psi} = g(\bar{\phi})$$
$$\Psi = G\Phi G^{\mathrm{T}}$$

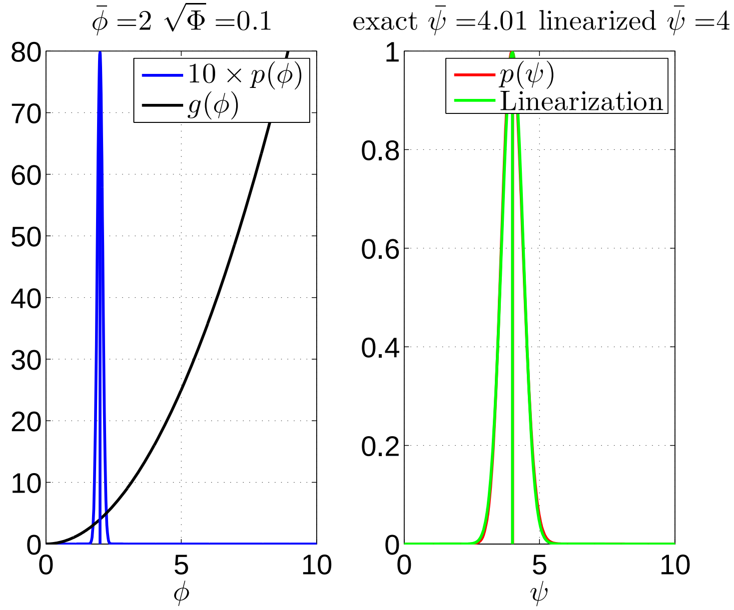
- With the linearization, the transformed mean is obtained by directly transforming the original mean.
- The covariance is obtained as in the linear transformation, where the transformation matrix is the Jacobian matrix of the nonlinear transformation.

Linearization Illustration

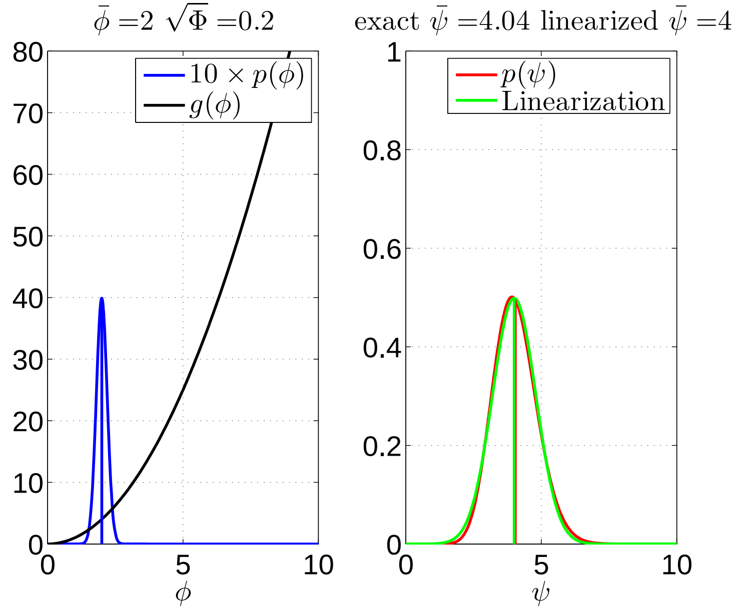
Example

- Let $\phi \sim \mathcal{N}(\phi, 2, \Phi)$ and $g(\phi) = \phi^2$.
- Change the standard deviation $\sqrt{\Phi} = 0.1, 0.2, \ldots, 1$.
- Observe the exact density $p(\psi)$ along with the approximation $\mathcal{N}(\psi; \bar{\psi}, \Psi)$ obtained from linearization.

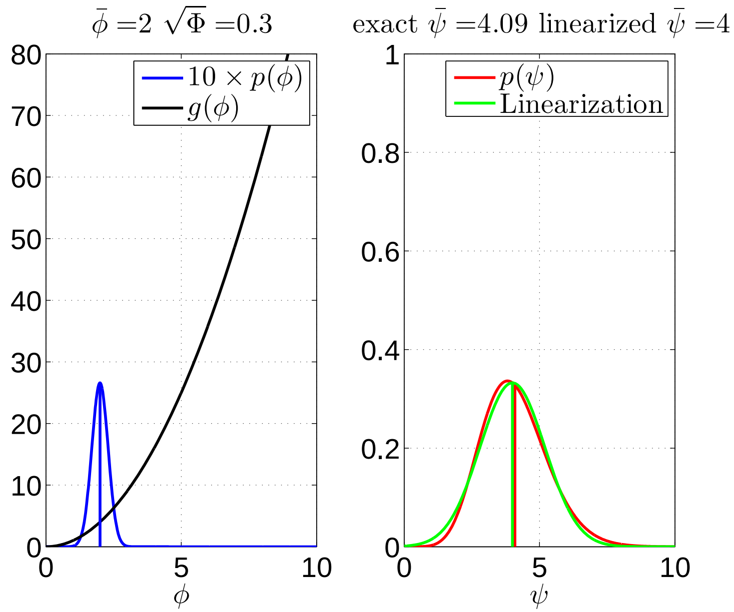
Linearization Illustration



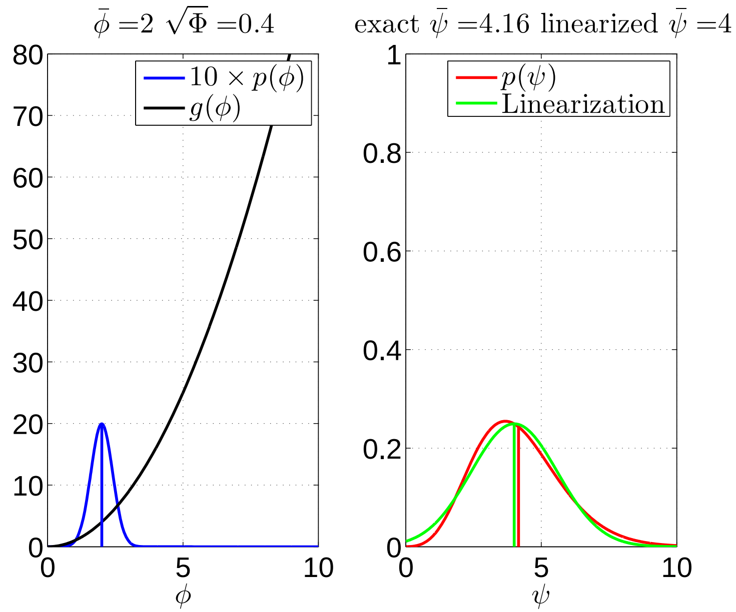
Linearization Illustration



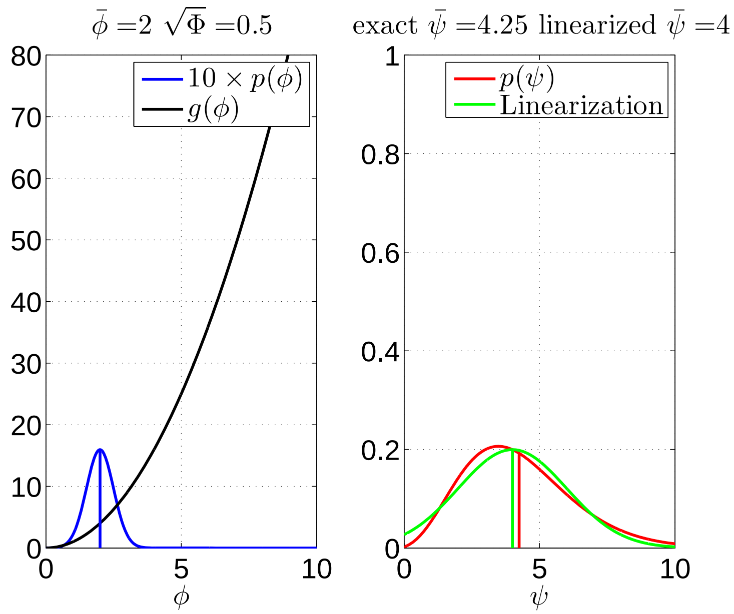
Linearization Illustration



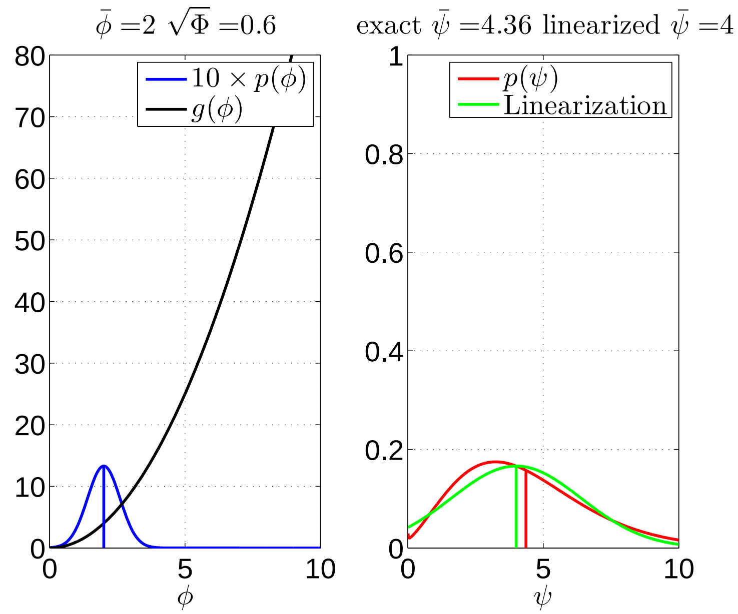
Linearization Illustration



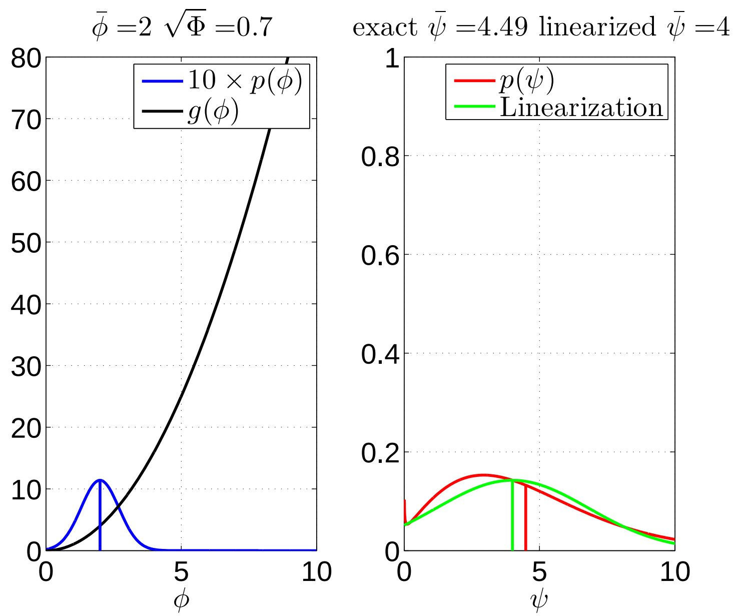
Linearization Illustration



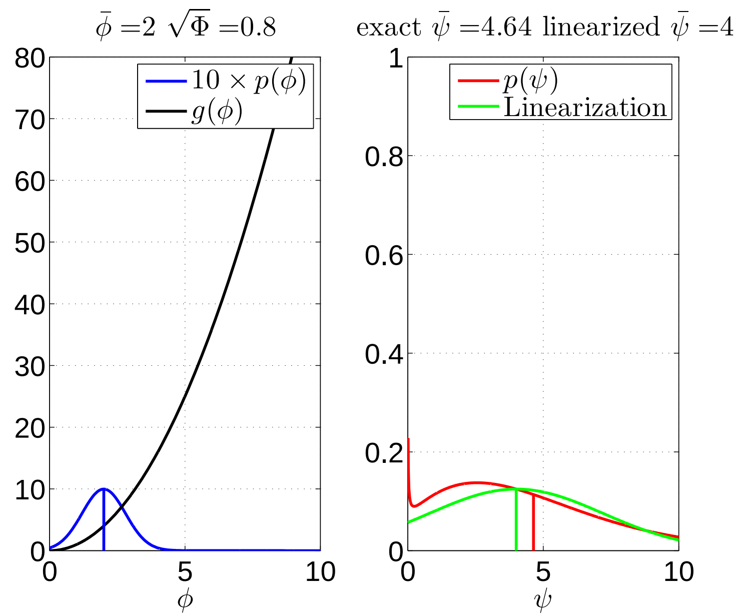
Linearization Illustration



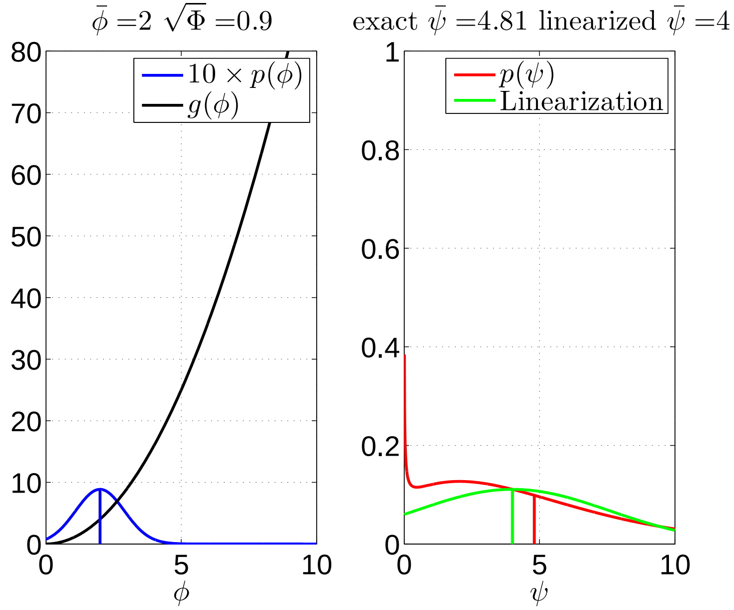
Linearization Illustration



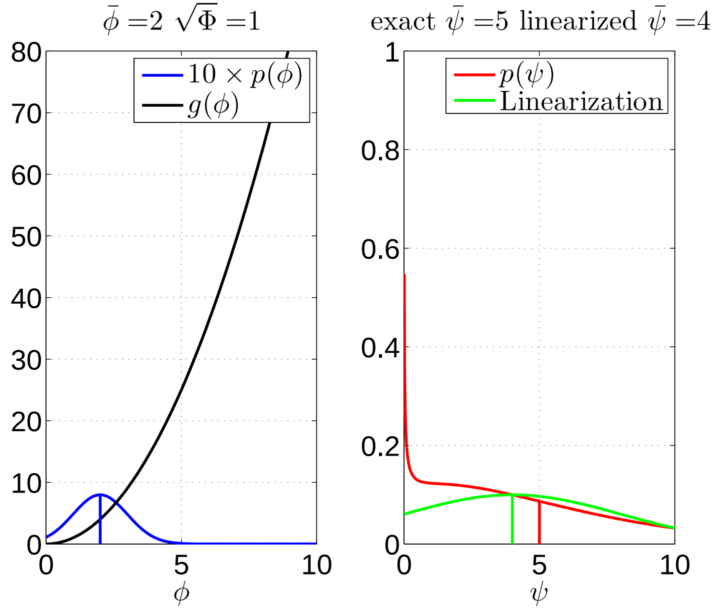
Linearization Illustration



Linearization Illustration



Linearization Illustration



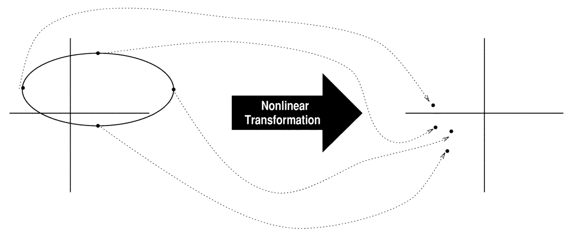
Linearization

- If the uncertainty is small in the variable to be transformed, the linearization gives good results.
- As the uncertainty grows, the performance of linearization degrades sometimes leading to terrible results.
- Linearization cares only about the information of transformation around the linearization point, hence it only works good locally. When the uncertainty grows, local results are bound to be bad.

Unscented Transform

- The second method of nonlinear transformation we are going to consider is the unscented transform.
- Unscented transform is based on using a number of points/particles (called as sigma-points) to represent the original Gaussian density $\mathcal{N}(\phi; \bar{\phi}, \Phi)$.

- The sigma-points are transformed with the nonlinear transformation $g(\phi)$.
- The mean and covariance of the transformed sigma-points give $\bar{\psi}$ and Ψ respectively.



The figure is taken from S.J. Julier, J.K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol.92, no.3, pp. 401–422, Mar. 2004.

Unscented Transform

Finding the Sigma-Points

We set the sigma-points and their weights for $\phi \sim \mathcal{N}(\phi, \bar{\phi}, \Phi)$ as

$$\begin{aligned} \phi^{(0)} &= \bar{\phi} & \pi^{(0)} &= \pi^{(0)} \\ \phi^{(i)} &= \bar{\phi} + \left[\sqrt{\frac{n_\phi}{1-\pi^{(0)}}\Phi}\right]_{:,i} & \pi^{(i)} &= \frac{1-\pi^{(0)}}{2n_\phi} \\ \phi^{(i+n_\phi)} &= \bar{\phi} - \left[\sqrt{\frac{n_\phi}{1-\pi^{(0)}}\Phi}\right]_{:,i} & \pi^{(i+n_\phi)} &= \frac{1-\pi^{(0)}}{2n_\phi} \end{aligned}$$

for $i = 1, \ldots, n_\phi$.

- Note that there are $2n_\phi + 1$ sigma-points.
- $\sqrt{\cdot}$ denotes the p.s.d. square-root of the matrix argument. `sqrtm(·)` or `cholcov(·)` in Matlab.
- $[\cdot]_{:,i}$ denotes the ith column of the matrix argument.

Unscented Transform

Finding the Sigma-Points

We set the sigma-points and their weights for $\phi \sim \mathcal{N}(\phi, \bar{\phi}, \Phi)$ as

$$\begin{aligned} \phi^{(0)} &= \bar{\phi} & \pi^{(0)} &= \pi^{(0)} \\ \phi^{(i)} &= \bar{\phi} + \left[\sqrt{\frac{n_\phi}{1-\pi^{(0)}}\Phi}\right]_{:,i} & \pi^{(i)} &= \frac{1-\pi^{(0)}}{2n_\phi} \\ \phi^{(i+n_\phi)} &= \bar{\phi} - \left[\sqrt{\frac{n_\phi}{1-\pi^{(0)}}\Phi}\right]_{:,i} & \pi^{(i+n_\phi)} &= \frac{1-\pi^{(0)}}{2n_\phi} \end{aligned}$$

for $i = 1, \ldots, n_\phi$.

- Note that we have $\sum_{i=0}^{2n_\phi} \pi^{(i)} = 1$ and

$$\sum_{i=0}^{2n_\phi} \pi^{(i)}\phi^{(i)} = \bar{\phi} \qquad \sum_{i=0}^{2n_\phi} \pi^{(i)}(\phi^{(i)} - \bar{\phi})(\phi^{(i)} - \bar{\phi})^{\mathrm{T}} = \Phi$$

Unscented Transform

Unscented Transform

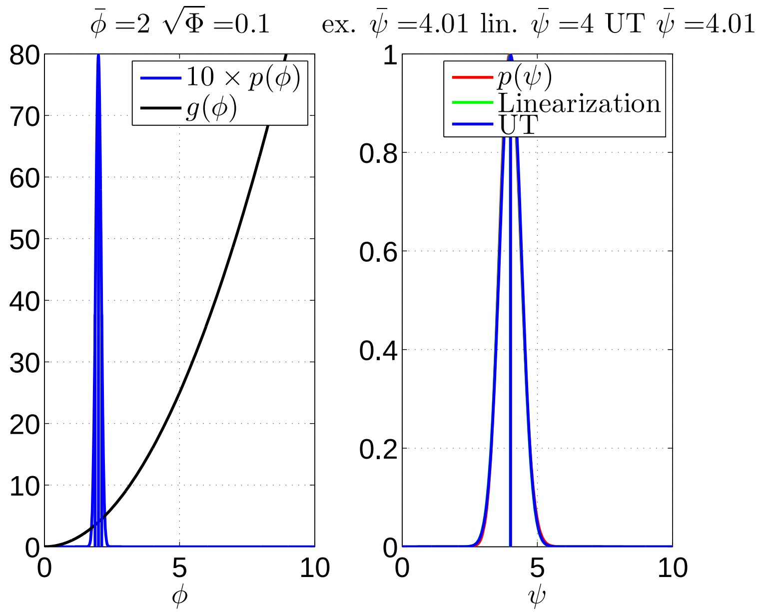
- Find the sigma-points and their weights $\{\pi^{(i)}, \phi^{(i)}\}_{i=0}^{2n_\phi}$.
- Transform the sigma-points with the transformation $g(\phi)$ as

$$\psi^{(i)} = g\left(\phi^{(i)}\right)$$

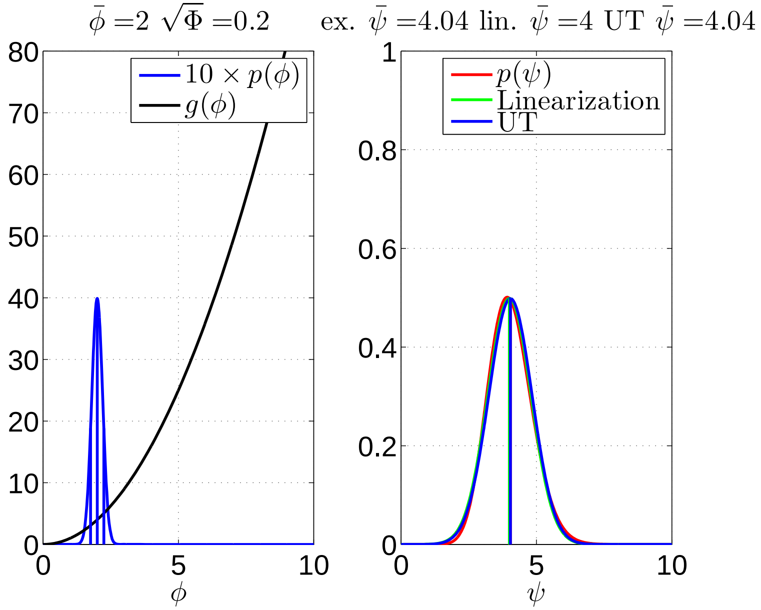
  for $i = 0, \ldots, 2n_\phi$.
- Find the transformed mean and covariance as

$$\bar{\psi} = \sum_{i=0}^{2n_\phi} \pi^{(i)}\psi^{(i)} \qquad \Psi = \sum_{i=0}^{2n_\phi} \pi^{(i)}(\psi^{(i)} - \bar{\psi})(\psi^{(i)} - \bar{\psi})^{\mathrm{T}}$$

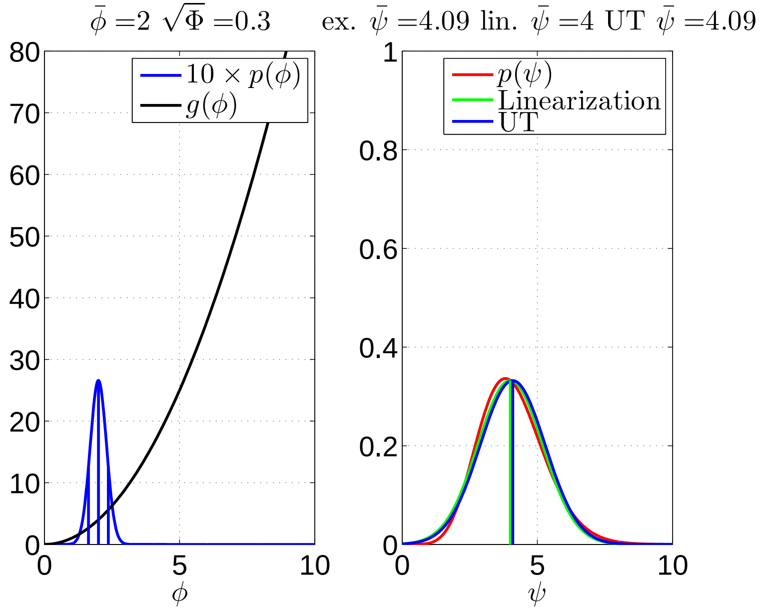
Unscented Transform Illustration



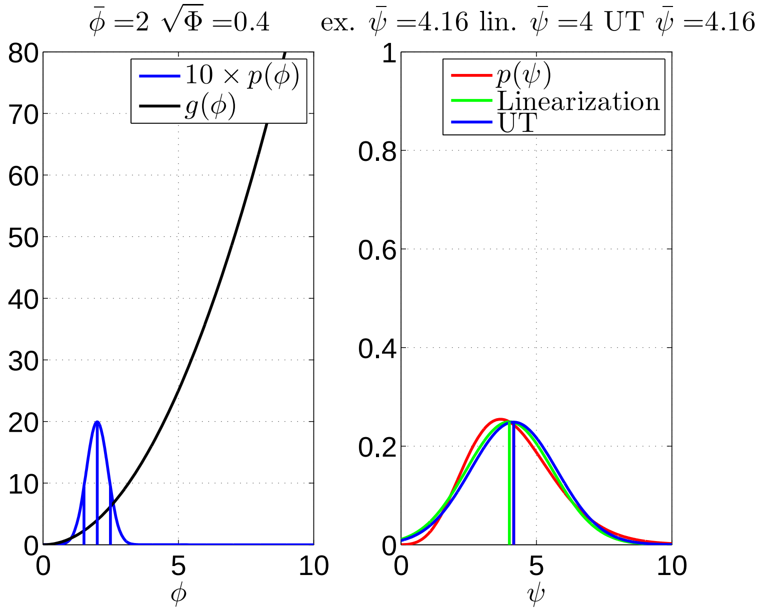
Unscented Transform Illustration



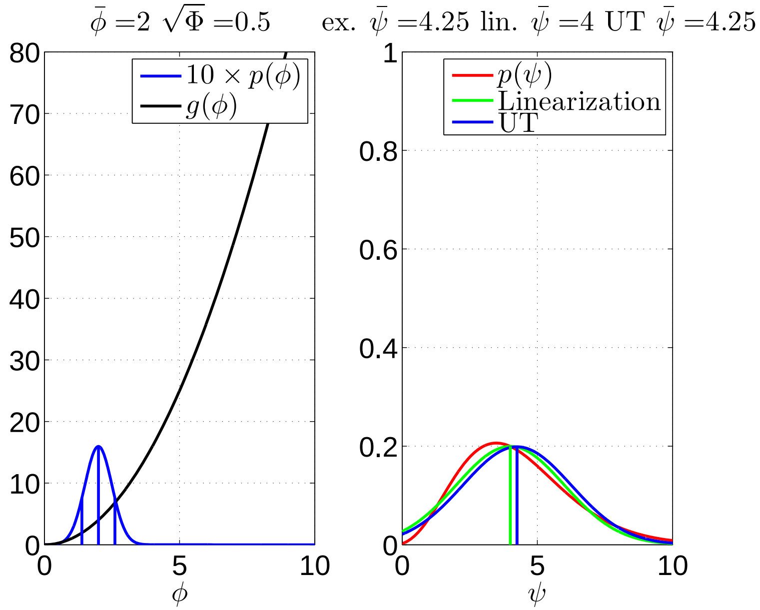
Unscented Transform Illustration



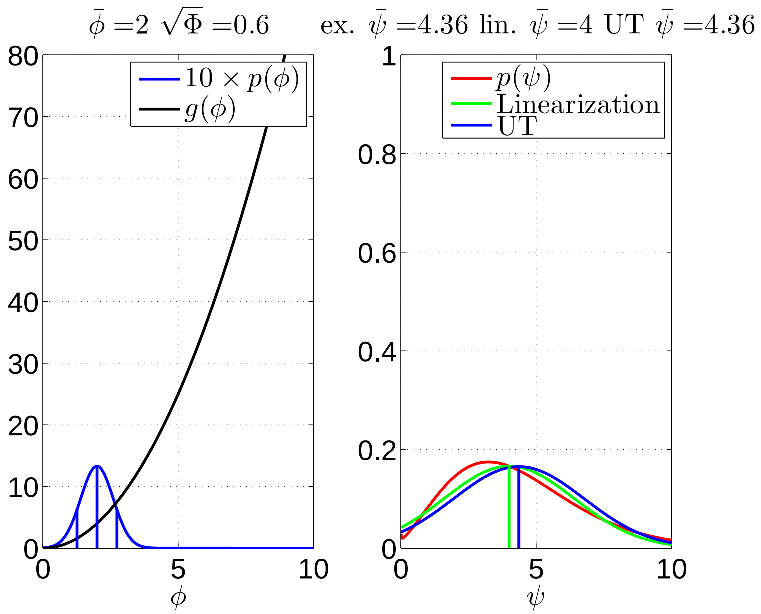
Unscented Transform Illustration



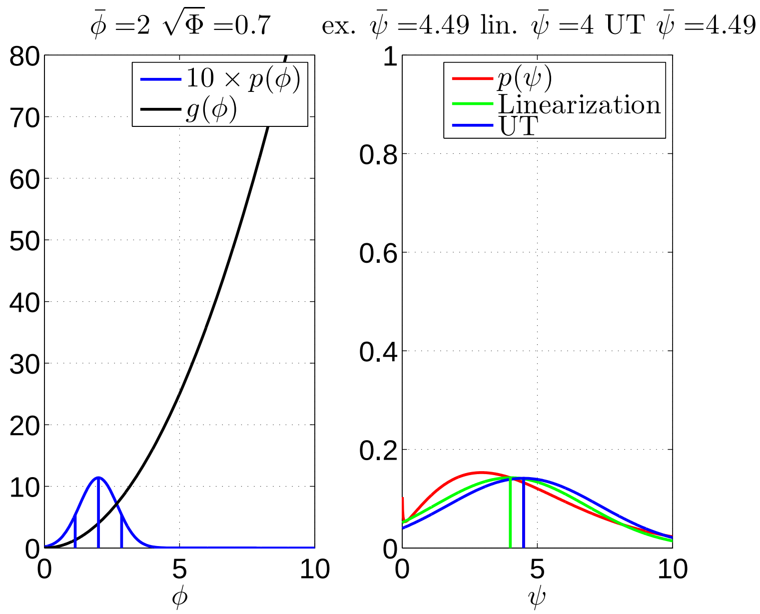
Unscented Transform Illustration



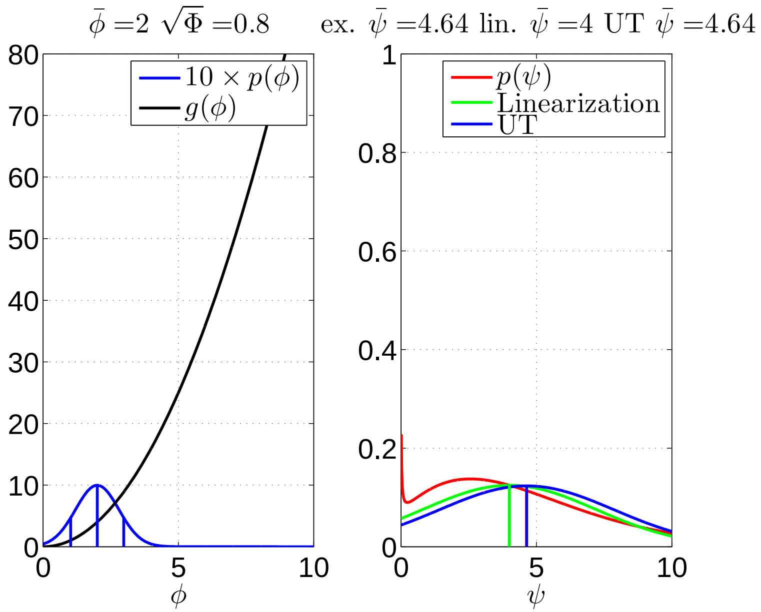
Unscented Transform Illustration



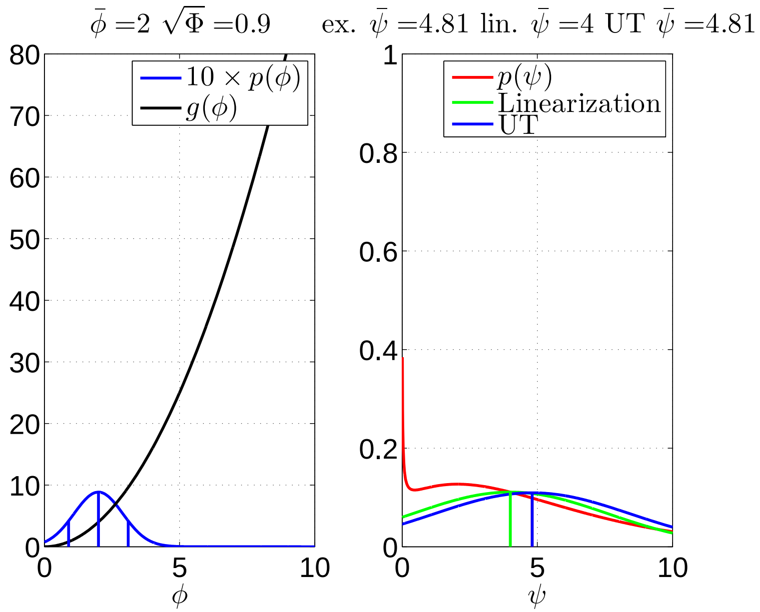
Unscented Transform Illustration



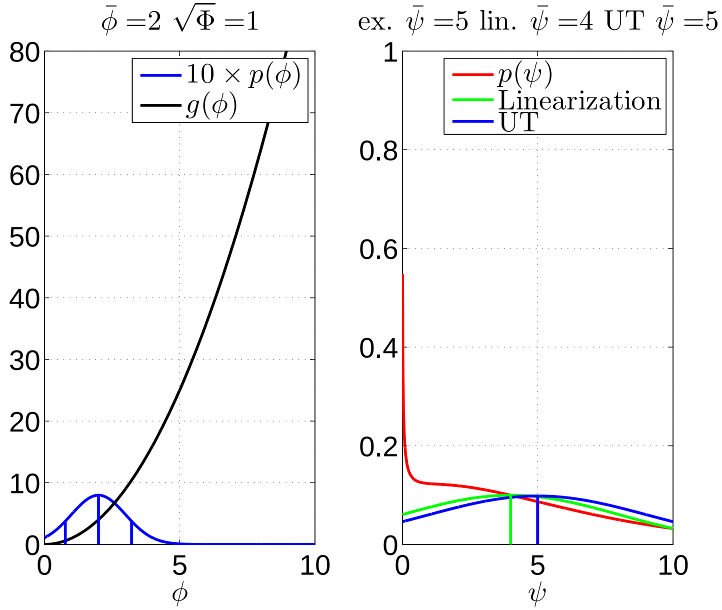
Unscented Transform Illustration



Unscented Transform Illustration



Unscented Transform Illustration



Back to Nonlinear and Non-Gaussian Bayesian State Estimation

Bayesian State Estimation

$$x_{k+1} = f(x_k) + w_k$$
$$y_k = h(x_k) + v_k$$

with $w_k \sim p(w_k)$, $v_k \sim p(v_k)$ and $x_0 \sim p(x_0)$.

- We can obtain a solution for the nonlinear non-Gaussian Bayesian state estimation problem using both linearization and unscented transform.
- These solutions are called extended Kalman filter (EKF) and unscented Kalman filter (UKF).
- For both approaches, we have to assume

$$p(w_k) \approx \mathcal{N}(w_k; 0, Q) \qquad p(x_0) \approx \mathcal{N}(x_0; \hat{x}_{0|0}, P_{0|0})$$
$$p(v_k) \approx \mathcal{N}(v_k; 0, R)$$

Extended Kalman Filtering

Extended Kalman Filter

- Start with $\hat{x}_{0|0}$, $P_{0|0}$, set $k = 1$.
- For each k
  - Prediction Update

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1})$$
$$P_{k|k-1} = F P_{k-1|k-1} F^{\mathrm{T}} + Q$$

where $F = \frac{\partial f}{\partial x_{k-1}}|_{x_{k-1}=\hat{x}_{k-1|k-1}}$.

  - Measurement Update

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1})$$
$$P_{k|k} = P_{k|k-1} - K_k S_{k|k-1} K_k^{\mathrm{T}}$$

where

$$\hat{y}_{k|k-1} = h(\hat{x}_{k|k-1}) \qquad S_{k|k-1} = H P_{k|k-1} H^{\mathrm{T}} + R$$
$$K_k = P_{k|k-1} H^{\mathrm{T}} S_{k|k-1}^{-1}$$

with $H = \frac{\partial h}{\partial x_k}|_{x_k=\hat{x}_{k|k-1}}$.

Unscented Kalman Filtering

Unscented Kalman Filter

- Start with $\hat{x}_{0|0}$, $P_{0|0}$, set $k = 1$.
- For each k
  - Prediction Update
    - Generate sigma-points and their weights $\{\pi^{(i)}, x^{(i)}_{k-1|k-1}\}_{i=0}^{2n_x}$ for $\mathcal{N}(x_{k-1}; \hat{x}_{k-1|k-1}, P_{k-1|k-1})$.
    - Transform the sigma-points.

$$x^{(i)}_{k|k-1} = f\left(x^{(i)}_{k-1|k-1}\right) \quad \text{for} \quad i = 0, \ldots 2n_x$$

    - Obtain the predicted state estimate $\hat{x}_{k|k-1}$ and its covariance $P_{k|k-1}$ as

$$\hat{x}_{k|k-1} = \sum_{i=0}^{2n_x} \pi^{(i)} x^{(i)}_{k|k-1}$$
$$P_{k|k-1} = \sum_{i=0}^{2n_x} \pi^{(i)} \left(x^{(i)}_{k|k-1} - \hat{x}_{k|k-1}\right)\left(x^{(i)}_{k|k-1} - \hat{x}_{k|k-1}\right)^{\mathrm{T}} + Q$$

Unscented Kalman Filtering

Unscented Kalman Filter

- Measurement Update
  - Generate sigma-points and their weights $\{\pi^{(i)}, x^{(i)}_{k|k-1}\}_{i=0}^{2n_x}$ for $\mathcal{N}(x_k; \hat{x}_{k|k-1}, P_{k|k-1})$.
  - Transform the sigma-points $\{\pi^{(i)}, x^{(i)}_{k|k-1}\}_{i=0}^{2n_x}$

$$y^{(i)}_{k|k-1} = h\left(x^{(i)}_{k|k-1}\right) \quad \text{for} \quad i = 0, \ldots 2n_x$$

  - Obtain the state estimate $\hat{x}_{k|k}$ and its covariance $P_{k|k}$ as

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1})$$
$$P_{k|k} = P_{k|k-1} - K_k S_{k|k-1} K_k^{\mathrm{T}}$$

where

$$\hat{y}_{k|k-1} = \sum_{i=0}^{2n_x} \pi^{(i)} y^{(i)}_{k|k-1} \qquad K_k = \Sigma_{xy} S^{-1}_{k|k-1}$$
$$S_{k|k-1} = \sum_{i=0}^{2n_x} \pi^{(i)} \left(y^{(i)}_{k|k-1} - \hat{y}_{k|k-1}\right)\left(y^{(i)}_{k|k-1} - \hat{y}_{k|k-1}\right)^{\mathrm{T}} + R$$
$$\Sigma_{xy} = \sum_{i=0}^{2n_x} \pi^{(i)} \left(x^{(i)}_{k|k-1} - \hat{x}_{k|k-1}\right)\left(y^{(i)}_{k|k-1} - \hat{y}_{k|k-1}\right)^{\mathrm{T}}$$

What is more?

- Extended and unscented Kalman filters are used extensively all over the world in many real applications.
- They are very useful when
  - Nonlinearities are mild.
  - Posterior densities are unimodal.
  - Uncertainties are small (i.e., SNR is high).
- When one or more of these conditions do not hold, they can
  - Simply give bad results.
  - Totally diverge.
- A more powerful framework that can be useful for such situations is particle filters.

Monte Carlo Methods

- The main idea is to approximate the posterior $p(x_k|y_{1:k})$ as

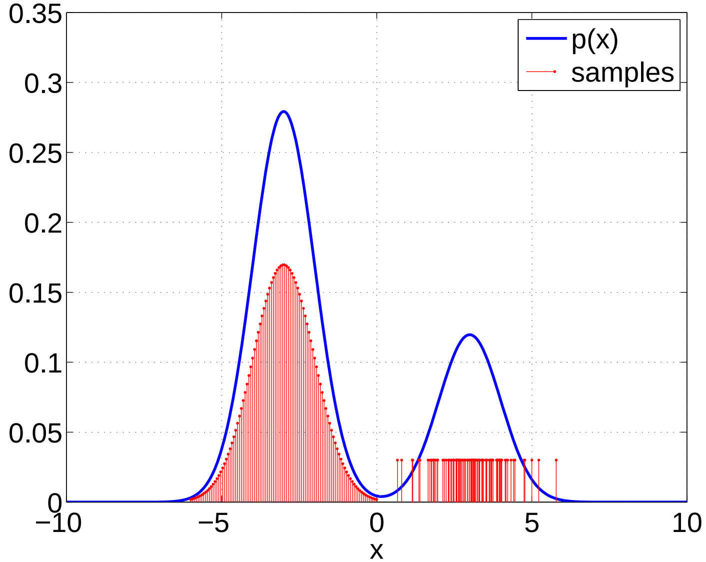
$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{N} \pi^{(i)}_{k|k} \delta_{x^{(i)}_{k|k}}(x_k)$$

where some state values $\{x^{(i)}_{k|k}\}_{i=1}^{N}$ called particles and weights $\{\pi^{(i)}_{k|k}\}_{i=1}^{N}$ are used.
- With Monte Carlo methods, taking any complicated integral simplifies to

$$\int g(x_k) p(x_k|y_{1:k}) \, \mathrm{d}x_k \approx \sum_{i=1}^{N} \pi^{(i)}_{k|k} g\left(x^{(i)}_{k|k}\right).$$

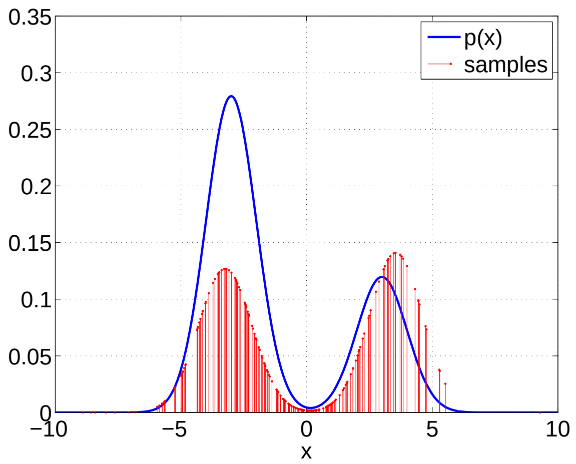
Example: Representation with Particles

$p(x) = 0.7\mathcal{N}(x; -3, 1) + 0.3\mathcal{N}(x; 3, 1)$ with 200 particles.



Example: Representation with Particles

- In general, both the weights and the proximity of the particles carry information.
- If the weight of a particle is high, one cannot directly conclude that density value is high there if the particle is in isolation.



Particle Filter

- Instead of propagating densities as

$$p(x_{k-1}|y_{1:k-1}) \xrightarrow{\text{prediction}} p(x_k|y_{1:k-1}) \xrightarrow{\text{update}} p(x_k|y_{1:k})$$

a particle filter propagates only the particles and the weights

$$\{\pi^{(i)}_{k-1|k-1}, x^{(i)}_{k-1|k-1}\}_{i=1}^N \xrightarrow{\text{prediction}} \{\pi^{(i)}_{k|k-1}, x^{(i)}_{k|k-1}\}_{i=1}^N \xrightarrow{\text{update}} \{\pi^{(i)}_{k|k}, x^{(i)}_{k|k}\}_{i=1}^N.$$

according to Bayesian density recursion.

- In some sense, a particle filter is a generalization of unscented Kalman filter to random particles instead of sigma-points.

Particle Filtering

Particle Filter

- Start with $x^{(i)}_{0|0} \sim p(x_0)$, $\pi^{(i)}_{0|0} = 1/N$ for $i = 1, \ldots, N$, set $k = 1$.
- For each k
  - Prediction Update
    - Sample process noise $w^{(i)}_{k-1} \sim p(w_{k-1})$.
    - Set the predicted particles and weights as

$$x^{(i)}_{k|k-1} = f\left(x^{(i)}_{k-1|k-1}\right) + w^{(i)}_{k-1} \qquad \pi^{(i)}_{k|k-1} = \pi^{(i)}_{k-1|k-1}$$

for $i = 1, \ldots, N$.

    - Obtain the predicted state estimate $\hat{x}_{k|k-1}$ and its covariance $P_{k|k-1}$ as

$$\hat{x}_{k|k-1} = \sum_{i=1}^N \pi^{(i)}_{k|k-1} x^{(i)}_{k|k-1}$$

$$P_{k|k-1} = \sum_{i=1}^N \pi^{(i)}_{k|k-1} \left(x^{(i)}_{k|k-1} - \hat{x}_{k|k-1}\right)\left(x^{(i)}_{k|k-1} - \hat{x}_{k|k-1}\right)^{\mathrm{T}}$$

Particle Filtering

Particle Filter

- Measurement Update
  - Set the estimated particles and weights as

$$x^{(i)}_{k|k} = x^{(i)}_{k|k-1}$$

$$\pi^{(i)}_{k|k} = \frac{\tilde{\pi}^{(i)}_{k|k}}{\sum_{i=1}^N \tilde{\pi}^{(i)}_{k|k}}$$

for $i = 1, \ldots, N$ where

$$\tilde{\pi}^{(i)}_{k|k} = \pi^{(i)}_{k|k-1} p\left(y_k \middle| x^{(i)}_{k|k}\right)$$

  - Obtain the state estimate $\hat{x}_{k|k}$ and its covariance $P_{k|k}$ as

$$\hat{x}_{k|k} = \sum_{i=1}^N \pi^{(i)}_{k|k} x^{(i)}_{k|k}$$

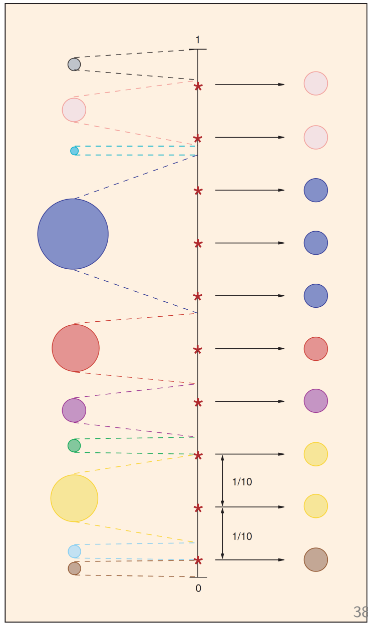
$$P_{k|k} = \sum_{i=1}^N \pi^{(i)}_{k|k} \left(x^{(i)}_{k|k} - \hat{x}_{k|k}\right)\left(x^{(i)}_{k|k} - \hat{x}_{k|k}\right)^{\mathrm{T}}$$

Particle Filtering

Particle Filter

- Resampling
  - A particle filter is useless without this step.
  - Without this step, all weights go to zero except one of them which becomes one.
  - This step removes the particles with negligible weights and replicates the particles with high weights.
  - Particle weights become all equal at the end of resampling.

Figure taken from P.M. Djuric, J.H. Kotecha, J. Zhang; Y. Huang; T. Ghirmai, M.F. Bugallo, J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol.20, no.5, pp. 19–38, Sep. 2003.



Particle Filtering: Resampling

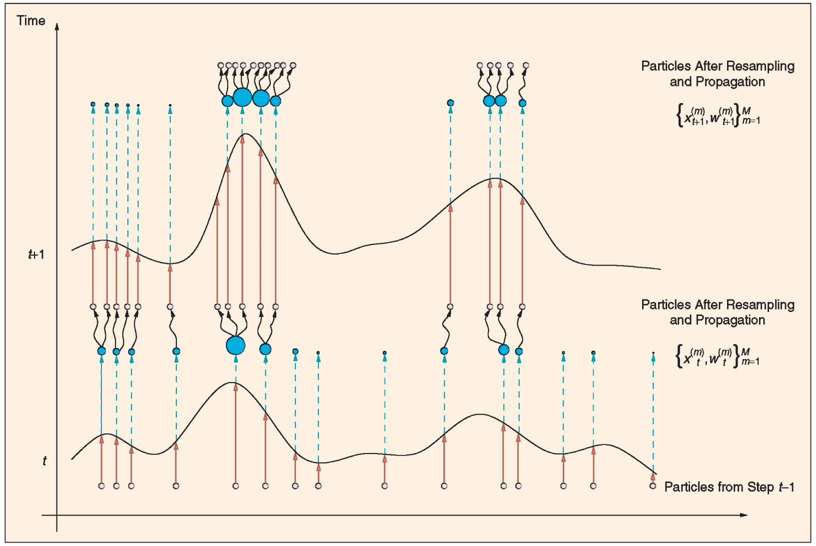


Figure taken from P.M. Djuric, J.H. Kotecha, J. Zhang; Y. Huang; T. Ghirmai, M.F. Bugallo, J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol.20, no.5, pp. 19–38, Sep. 2003.

References

KF and EKF

- B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, 1979.
- A. Gelb, *Applied Optimal Estimation*, MIT Press, 1974.
- A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, 2001.

UKF

- S.J. Julier, J.K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol.92, no.3, pp. 401–422, Mar. 2004.
- S.J. Julier, J.K. Uhlmann, "Data Fusion in Nonlinear Systems", Ch. 15, in *Handbook of Multisensor Data Fusion: Theory and Practice*, M. E. Liggins, D. Hall, J. Llinas (Editors), Taylor & Francis, 2009.

References

PF

- M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol.50, no.2, pp.174–188, Feb. 2002.
- F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Transactions on Signal Processing*, vol.50, no.2, pp.425–437, Feb. 2002.
- P. M. Djuric, J. H. Kotecha, J. Zhang; Y. Huang; T. Ghirmai, M. F. Bugallo, J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol.20, no.5, pp. 19–38, Sep. 2003.
- A. Doucet and X. Wang, "Monte Carlo methods for signal processing: a review in the statistical signal processing context," *IEEE Signal Processing Magazine*, vol.22, no.6, pp. 152–170, Nov. 2005.
- F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerospace and Electronic Systems Magazine*, vol.25, no.7, pp.53–82, July 2010.